Pattern-based decompositions for human resource planning in home health care services

Yalçindağ, Semih; Cappanera, Paola; Grazia Scutellà, Maria; Şahin, Evren; Matta, Andrea

Pattern-Based Decompositions for Human Resource Planning in Home Health Care Services

Semih Yalçındağ

Industrial and Systems Engineering Department
Yeditepe University, Istanbul, Turkey
(email: semih.yalcindag@yeditepe.edu.tr)

Paola Cappanera

Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Firenze
(email: paola.cappanera@unifi.it)

Maria Grazia Scutellà

Dipartimento di Informatica
Università di Pisa
(email: scut@di.unipi.it)

Evren Şahin

Laboratoire Génie Industriel
CentraleSupelec
(email: evren.sahin@centralesupelec.fr)

Andrea Matta

Department of Industrial Engineering and Management
Shanghai Jiao Tong University,
(email: matta@sjtu.edu.cn)

March 4, 2016

1

Abstract

Home health care services play a crucial role in reducing the hospitalization costs due to the increase of chronic diseases of elderly people. At the same time, they allow to improve the quality of life for those patients that receive treatments at their home. Optimization tools are therefore necessary to plan service delivery at patients' homes. Recently, solution methods that jointly address the assignment of the patient to the caregiver (assignment), the definition of the days (pattern) in which caregivers visit the assigned patients (scheduling), and the sequence of visits for each caregiver (routing) have been proposed in the scientific literature. However, the joint consideration of these three levels of decisions may be not affordable for large providers, due to the required computational time.

In order to combine the strength and the flexibility guaranteed by a joint assignment, scheduling and routing solution approach with the computational efficiency required for large providers, in this study we propose a new family of two-phase methods that decompose the joint approach by incrementally incorporating some decisions into the first phase. The concept of pattern is crucial to perform such a decomposition in a clever way. Several scenarios are analyzed by changing the way in which resource skills are managed and the optimization criteria adopted to guide the provider decisions. The proposed methods are tested on realistic instances. The numerical experiments help to identify the combinations of decomposition techniques, skill management policies and optimization criteria that best fit with problem instances of different size.

Keywords: Home Health Care; Mathematical Programming; Optimization; Skill Management

1 Introduction

Human resource planning defines, over a given time horizon, the composition of the required workforce to meet the service goals of the system under study. In the literature, this planning process is classified under different categories: permanence centered planning, fluctuation centered planning, mobility centered planning and project centered planning (Causmaecker et al., 2005). This paper focuses on the short term planning process for mobility centered planning (Brucker et al., 2011), where human resources (operators) travel to perform on-site service activities at customer places. For such services, innovative decision making methods are required to support the planning process.

Home Health Care (HHC) is a relevant example of such services that emerge as an increasingly promising alternative for providing health and social services

to patients cared at home (Kucukyazici et al., 2011; Matta et al., 2014; Sahin and Matta, 2014). Many factors drive the demand for HHC. Among these, we can cite the actual demographic trends, the changes in the epidemiological landscape of diseases and the availability of new support technologies. In the HHC service process, that involves several types of operators (e.g., nurses, physicians, social workers, home support workers, etc.), the human resource planning process is of particular interest. As such, the human resource planning process for Home Health Care (HRPHHC) consists of several decisions such as capacity planning, partitioning of the territory where the HHC service operates into districts, allocation of resources to districts, assignment of operators to patients (or to visits), scheduling and routing of operators (Matta et al., 2014; Sahin and Matta, 2014).

In this paper we focus on the last three levels of planning that are the assignment, scheduling and routing decisions. Given a planning horizon, a set of patients where each patient has a specific care plan (i.e. weekly care service requests asking for specific operator qualifications/skills), and a set of operators characterized by skills, the addressed HRPHHC problem asks to (i) assign the operators to the patients by taking into account the compatibility between requests and operator skills (assignment decisions), (ii) schedule the patient requests during the planning horizon (scheduling decisions), and (iii) determine the tour each operator has to perform on each day of the planning horizon (routing decisions). We consider variously skilled operators, various patient care requests, multiple planning period (which is usually a week) and continuity of care constraints (i.e., the assignment patient-operator is not changed during the whole planning horizon (Lanzarone et al., 2012)).

In the state-of-the-art literature, the HRPHHC problem is usually solved in cascade: first operators are assigned to patients on a geographical basis so as to match the skills demanded by patient care requirements with skills owned by each operator; second, the routing of each operator is determined. Such *two-phase methods* generally deal with a daily planning horizon. In a more flexible way, the problem can be solved without decomposing the three levels of planning decisions, but jointly addressing assignment, scheduling and routing decisions over the given planning horizon. We refer to this as the *single-phase method*.

An innovative modelling device to combine the three levels of decisions has recently been proposed, where services are offered according to a set of a priori given *patterns*, i.e., possible schedules for patients' requests (Cappanera and Scutellà, 2015). It has been shown that, by properly selecting the pattern generation mechanism, the single-phase method is able to compute very good quality solutions, near to the optimal ones. Large amounts of time and memory resources may be required, however, in the case of large instances. Hence, the use of the pattern device coupled with a two-phase decomposition approach seems to be suitable for solving large scale HRPHHC instances by combining efficacy (solution

quality) with efficiency (solution time). This represents the subject of this paper.

Precisely, the main contribution of this paper is the proposal of a new family of two-phase methods, based on patterns, where assignment, scheduling and routing decisions over the considered planning horizon are taken in two steps, as in the traditional two-phase methods. Such steps, however, are properly coordinated by means of the pattern mechanism introduced in (Cappanera and Scutellà, 2015). In this way the strength of the single-phase method is combined with the computational efficiency of the traditional two-phase approaches.

The pattern-based two-phase methods are investigated also computationally, by providing a comparative study among them. Specifically, since the proposed methods vary in terms of flexibility and efficiency, the goal is to identify the most effective and efficient ones. In addition, alternative objective functions based on operating cost minimization and on social equity criteria (i.e., balancing operator workloads) are compared under different conditions of skill management. The presented methods are indeed relaxations of the exact, single-phase approach in (Cappanera and Scutellà, 2015). Therefore, such an exact approach is considered in order to evaluate the solution quality guaranteed by the pattern-based two-phase methods.

Numerical analysis is carried out on realistic problem instances. Results show that some of the proposed two-phase methods provide a powerful tool to solve HRPHHC, leading to high quality solutions in a short computational time. In other words, they appear to be a very promising tool to address HRPHHC with both efficacy and efficiency, especially for large size instances. This finding is significantly important in practice where many HHC providers have to deal with large scale problems, characterized by several hundreds of patients located in a single district.

The paper is organized as follows. Section 2 provides a literature review on the short term HRPHHC. Section 3 describes the problem and also provides an overview of the methods presented in this work. Such methods are then detailed in Section 4. Results from an extensive experimentation are reported in Section 5. Finally, conclusions are drawn in the last section.

2 Literature Review

The literature related to assignment, scheduling and routing problems in HHC services has been reviewed by two recent works (Hulshof et al., 2012; Yalcindag et al., 2012). Hulshof et al. (2012) propose a taxonomic review on planning-related decisions in health care services, including HHC. Yalcindag et al. (2012) review the assignment and routing problems in HHC. Among the existing works already mentioned in these papers, this section summarizes and classifies the most

relevant ones. Papers are thus grouped based on the length of the planning period considered, the objective function used as well as constraints imposed. This enables to point out the main differences of the problems studied in this paper with respect to those addressed in the literature. The solution approaches proposed by authors are also briefly presented.

Most of existing literature is devoted to the daily HHC planning problems. Among these, Cheng and Rich (1998) develop a daily scheduling problem as a multi-depot Vehicle Routing Problem (VRP) with time windows (MDVRPTW) that considers compatibility information between patients and operators. The problem is formulated as a mixed integer linear program. The objective is to minimize the total cost associated with the amount of overtime hours of full-time and part-time nurses. Meanwhile, this objective is pursued with respect to constraints such as visiting each patient exactly once, assigning each nurse to at least one patient, starting and ending at the operator's home, taking a lunch break within the given time interval and respecting the constraint of maximum nurse shift length. Since the problem is a complex combinatorial optimization problem, Cheng and Rich (1998) develop a two-phase heuristic. The first phase falls into the parallel tour-building procedure category. The identified tours are then improved in the second phase using local search by adjusting assignments and inserting omitted patients.

Eveborn et al. (2006) formulate the scheduling problem as a VRPTW with the set partitioning model and solve it heuristically by using a repeated matching algorithm. The objective is to minimize a total cost related to the travel time, scheduled hours, preferences, etc., while respecting the following constraints: time windows for visits, operators' skill requirements, and accomplishment of each visit by one operator.

Thomsen (2006) addresses the daily scheduling problem as a VRPTW with shared visits (i.e., visits requiring two operators). The objective of this model is to minimize the total travelling cost and the number of visits that are carried out by non reference care givers. The constraints of the model are as follows: respecting the visits' and operators' time windows, assigning at least one visit to each operator and starting/ending a shared visit at the same time. The model is solved by using a new insertion method that forms an initial solution for a variant of tabu search.

Akjiratikarl et al. (2007) generate daily schedules by using the VRPTW. They minimize the total distance travelled with respect to visits' and operators' time windows and assignment of each visit to only one operator. The solution procedure incorporates the Local Improvement Procedure into the Particle Swarm Optimization technique to improve the identified solutions. The initial solutions are generated using the Earliest Start Time Priority with the Minimum Distance Assignment technique.

5

Bredstrom and Ronnqvist (2008) develop a mathematical model that integrates synchronization and precedence constraints between visits by extending the traditional VRP formulation. They use a heuristic approach based on the local branching heuristic to solve their model. In their previous study (Bredstrom and Ronnqvist, 2007), they develop a branch-and-price algorithm to solve the same model without including the precedence constraints.

Trautsamwieser et al. (2011) propose a model for securing the HHC services when natural disasters appear. They develop the daily scheduling model as a VRP with state-dependent breaks. The objective of the model is to minimize the sum of travel times and waiting times as well as the dissatisfaction levels of patients and operators. The proposed model is first solved for small scale data with a state of art solver, and then for real life-sized data with a neighborhood search based heuristic.

Rasmussen et al. (2012) address the daily scheduling problem as a multi-depot VRPTW and temporal connections between visits. They use the same formulation as in Bredstrom and Ronnqvist (2007) with some differences. Visits can be uncovered and a multi-criteria objective is adopted to minimize the number of uncovered visits with the total distance traveling costs while maximizing the operator-visit preferences. In particular, in the objective function they assign a higher priority to the uncovered visit criterion than the other criteria. Constraints of this model include: each visit can be covered exactly once or left as uncovered, operators can only handle allowed visits, visits' and operators' time windows and precedence relations of visits.

Despite the fact that HHC providers often make decisions on a multi-period planning horizon, only few works focus on the weekly planning problem of HHC services (Begur et al., 1997; Nickel et al., 2012; Borsani et al., 2006; Cappanera and Scutellà, 2015; Gamst and Jensen, 2012; Cattafi et al., 2012; Trautsamwieser and Hirsch, 2014). Begur et al. (1997) propose a Spatial Decision Support System (SDSS) that contains a special module for the daily scheduling of operators' activities. This module assigns simultaneously operators to visits and generates the sequence in which the visits should be executed. The tool is based on a heuristic approach to minimize the total travelling time while respecting constraints related to the route construction, operators time windows, and skills requirements.

Borsani et al. (2006) propose a multi-objective formulation that includes the number of covered visits, continuity of care constraints and travel costs. Patients' and operators' preferences are modeled as soft constraints, but routing decisions are not addressed. A state-of-the-art solver is used for the two mathematical formulations.

Nickel et al. (2012) address the weekly scheduling problem as the combination of the VRPTW and the nurse rostering problem. The objective of the proposed model is minimizing the weighted sum of the patient-nurse loyalty (conti-

6

nuity of care), unscheduled tasks, the overtime costs, and the traveling distance. As solution approach, they use different meta-heuristics combined with methods from constraint programming that allows a very flexible treatment of realistic constraints.

Gamst and Jensen (2012) propose a mathematical model and a Branch-and-Price solution approach. They deal with regular visiting times at the patients by the same group of operators. Visiting time windows and working time windows of the operators are not considered as hard constraints but are handled within a weighted objective function.

Cattafi et al. (2012) propose a solution approach based on constraint programming where they assign operators to the patients and solve a traveling salesman problem for each operator on each day. Time windows are not addressed and the operators may work also in intramural health care facilities. The total workload of the operators and the number of different operators visiting a client are minimized.

Trautsamwieser and Hirsch (2014) study the case where patients need to be visited one or several times during the week by appropriately skilled operators. Visits have predefined time windows. Additionally, working time requirements for the operators such as breaks, maximum working time per day, and daily as well as weekly rest times are considered. The authors propose a Branch-Price-and-Cut solution approach to solve the problem exactly, using the solutions of a variable neighborhood search solution approach as upper bounds. Their algorithm is capable of solving to optimality real-life based test instances with up to nine nurses, 45 clients, and 203 visits during the week.

Lastly, Cappanera and Scutellà (2015) present a single-phase method to the weekly planning of HHC services, where assignment, scheduling and routing decisions are jointly addressed via patterns. In fact, a pattern based mechanism is proposed to combine the diverse levels of decisions. Furthermore, skill compatibility between patients and operators is addressed. Specifically, the authors develop a hierarchical skill management policy where an operator with a specific skill is allowed to serve patients requiring lower level skill visits. The mathematical model aims at balancing the workload among operators.

The two-phase methods presented in this paper originate as a decomposition of the single-phase method in Cappanera and Scutellà (2015), with the aim to achieve more efficiency still guaranteeing the quality of the solutions found by an exact approach. Assignment, scheduling and routing decisions are taken in two steps, by defining two mathematical programming models related one each other to keep the consistency of the problem. Assignment decisions are always taken in the first phase. However, differently than in traditional two-phase approaches, in some cases also scheduling decisions are addressed in the first phase, and this is possible by suitably exploiting the pattern mechanism mentioned above, which is now an instrument to coordinate the two phases of the approach.

7

3   Problem description

This section has two objectives. First, Section 3.1 presents the general assumptions we use in the HRPHHC problem studied in this paper. Then, Section 3.2 presents the methods used in this study.

3.1   Assumptions and notation

Given a planning horizon $\mathcal{H}$, usually a week, a set of patients with an associated *care plan*, i.e. weekly requests that need specific skills to be operated, and a set of operators where each operator is characterized by a specific skill, the analyzed problem asks to assign the operators to the patients by taking into account the compatibility between the requests and operator skills (*assignment decisions*), schedule the patient requests during the planning horizon (*scheduling decisions*), and determine the tour each operator has to perform on each day of the planning horizon (*routing decisions*).

In a more formal way, the problem can be defined on a complete directed network $G = (\mathcal{N}, \mathcal{A})$ having n nodes in the set $\mathcal{N} = \{1, \ldots, n\}$, where each node j corresponds to a patient. We assume to have an extra node (node 0), which is used to denote the basis of operators, i.e. the daily tour of each operator starts/ends from/at node 0. The set $\mathcal{A}$ represents the possible trajectories among the nodes of the network. Arcs are labeled with $t_{jj'}((j, j') \in \mathcal{A})$ representing the distances among the nodes.

A set $\mathcal{K} = \{1, \ldots, K\}$ of K levels of skill is considered for either patients (each patient skill represents a set of specific care requirements) or operators (each operator skill represents a set of specific competencies), where skill K corresponds to the most complete competency level and skill 1 to the least complete one. As an example, if $K = 2$, skill 1 would refer to ordinary requests or basic operator competencies, whereas skill 2 can correspond to palliative requests or advanced operator competencies.

A *care plan* r_j is assumed to be known for each patient j. Indeed, the vector r_j specifies, for each level of skill, the number of visits required by patient j in the whole planning horizon $\mathcal{H}$ relatively to that skill. Hence, according to this, each care plan r_j has K components denoted with r_{jk} (with $k \in \mathcal{K}$) that represent the number of visits of skill k required by patient j in the planning horizon.

The set $\mathcal{O}$ represents the (skilled) operators available in the planning horizon, whereas subsets $\mathcal{O}_d \subseteq \mathcal{O}$ represent the operators available on a particular day d, for each $d \in \mathcal{H} = \{1, \ldots, H\}$. A "hierarchical structure of skill levels" refers to the situation where an operator with skill k can be assigned to all care requests characterized by a skill level inferior or equal to k. We denote with $\mathcal{S}_k \subseteq \mathcal{O}$ the subset of operators having at least skill k.

The HRPHHC decisions are taken by using the *pattern* concept. As such, care requests associated with each patient are satisfied by choosing the most appropriate pattern within a set $\mathscr{P}$ of a priori given patterns. For example, if a patient requires three visits of a given skill in a week, visits can be operated according to the pattern Monday-Wednesday-Friday or Monday-Tuesday-Thursday. Formally, for each pattern $p \in \mathscr{P}$ we define $p(d) = 0$ if no service is delivered at day d, while it is $p(d) = k$ if a visit of skill k is operated according to pattern p on day d. The implicit modeling assumption is that a patient may receive at most one visit per day. This assumption excludes hospitalization-at-home service, which requires intensive care services, and acute patients in general. In case more visits can be delivered in a single day, the pattern definition can be extended to consider this (Cappanera and Scutellà (2015)).

The addressed HRPHHC problem thus consists in: i) assigning one operator to each patient j (*assignment* decisions), ii) assigning a pattern p from $\mathscr{P}$ to each patient j in $\mathscr{N}$, so scheduling the requests of j expressed by r_j during the planning horizon (*scheduling* decisions), and iii) computing the tour of each operator for each scheduled day (*routing* decisions). In addressing these three levels of decisions, skill constraints (i.e. the compatibility between requirements associated with patient requests and skills of operators) as well as daily workload constraints for operators are taken into account. *Continuity of care* constraint (i.e., a unique operator has to be assigned to each patient over $\mathscr{H}$) is included. Since service providers may by guided by several criteria in the HRPHHC, alternative objective functions are considered, as better specified in the next section.

It is important to emphasize that many significant differences do exist in the planning of home care organisation operations, which induce the development of different models and methods. An analysis of the main types of organisations is depicted in Matta et al. (2014), where a distinction among public service providers, and private non for profit/for profit providers is outlined depending on their corporate status. Another classification of home care providers is related to the pathologies suffered by their patients and the patient characteristics.

The HRPHHC problem studied in this paper refers to providers which are mainly dedicated to palliative care and terminal patients. This is the most common case in Italy as stated in Cappanera and Scutellà (2015). However its characteristics are very general; in fact, most of them constitute a common kernel to other types of health care service organisations. In contrast, the issues that characterize this HRPHHC problem are mainly the low importance of time windows constraints, the care continuity management and a hierarchical structure of the skills associated with patients and caregivers. In particular, although time windows imposed on patient's visits can be crucial in several home care scenarios, especially for patients requiring a timely administration of drugs, in the palliative

context the satisfaction of time windows indicated by the patients is typically not pursued. Furthermore, another issue that characterizes the palliative context is that patient's visits and operators are associated with levels of skill that are managed hierarchically. This opportunity is perceived positively by the patients since it increases their confidence in the service they receive.

The issues that feature the palliative context, however, do not narrow the applicability of the methods we propose to different home care contexts. Extensions can be easily designed that may help in coping with the request to satisfy time windows indicated by the patients and those relative to operator shifts, like a morning and an afternoon shift in each day, and with scenarios where skills are not organized hierarchically (Cappanera and Scutellà, 2015). Work time constraints such as break times, minimum rest times between shifts and minimum weekly rest times, as well as synchronization constraints among visits at the same patients, not addressed in this paper, also deserve consideration. They will be the subject of future work.

3.2 Overview of methods

This section presents four alternative pattern-based methods to tackle the HRPHHC problem stated before, as well as some relevant variants. One is the single-phase method in (Cappanera and Scutellà, 2015), while the others are two-phase methods originating from it. As previously introduced, such two-phase methods are obtained by suitably decomposing the single-phase model in (Cappanera and Scutellà, 2015) to enhance its computational efficiency. Precisely, each two-phase method is defined by two mathematical programming models related to each other to keep the consistency of the problem. The methods differ by their degree of flexibility, where flexibility refers to the possibility of exploring a wider solution space: the more a method is flexible, the wider its solution space which corresponds to postpone the decisions in the last phase. With the less flexible method assignment decisions are taken in the first phase (thus guiding the other remaining decisions). With the most flexible method no decision has a preeminent role with respect to the others and all of them are taken simultaneously in one shot (this is the single-phase method).

In the less flexible method, patients are assigned to operators by looking only at the overall load of their care plan and at the availability of operators over the entire programming period. Such operator-patient assignments may eventually generate infeasibilities in the second phase, where scheduling and routing decisions are taken for each day of the planning period. In an attempt to limit these infeasibilities, we propose two intermediate methods characterized by different levels of flexibility. They both integrate scheduling decisions in the first phase.

In the less flexible method of the two intermediate ones, scheduling decisions

10

are taken only in the first phase in addition to the assignment decisions. Then, once each patient has been assigned to one operator (assignment) and once the day pertaining to each visit specified in the care plan of the patients has been fixed (scheduling), the tour each operator has to perform every day of the planning horizon is consequently optimized in the second phase.

In the more flexible method of the two intermediate ones, the assignment is guided by a tentative schedule of visits, but such a schedule is not fixed in the second phase. Indeed, scheduling decisions are re-considered in the second phase together with routing decisions. Integrating the scheduling decisions into the first phase has the aim of better guiding assignment decisions.

A graphical representation of the four methods described above is given in Figure 1, by using the notation introduced in Table 1. Specifically, in Table 1 the main decisions are reported together with their description, the label used in the following to refer them and the name of the variable that will be used in the models definition to represent the related decision. In particular, u_{ij} is a binary variable set to one when patient j is assigned to operator i; z_{jp} is a binary variable set to one if patient j is served according to pattern p while $x_{jj'}^{id}$ is a binary variable set to one when the tour of operator i on day d includes consecutively patients j and j'. Figure 1 shows how each of the four methods ranks with respect to flexibility (x-axis) and in which level the decisions are taken (y-axis).

With regards to decisions taken, each of the four methods is identified by a two field string where each field contains the label(s) of decisions taken. As an example, the string $A|S+R$ denotes the rigid method where assignment decisions characterize the first phase while scheduling and routing decisions are taken in the second phase. Such a notation will be enriched in the next paragraphs with the information coming from the objective function used to guide each phase.

Table 1: Type of decisions

Decision	Label	Description	Variable
Assignment	A	assign a set of patients to each operator	u_{ij}
Scheduling	S	fix the day of each visit	z_{jp}
Routing	R	sequence the visits in each day for each operator	$x_{jj'}^{id}$

An interesting feature of two-phase methods is that different objective functions may be selected to guide the optimization at each phase. In other words, these methods are inherently bicriteria. In our study, the first phase of each
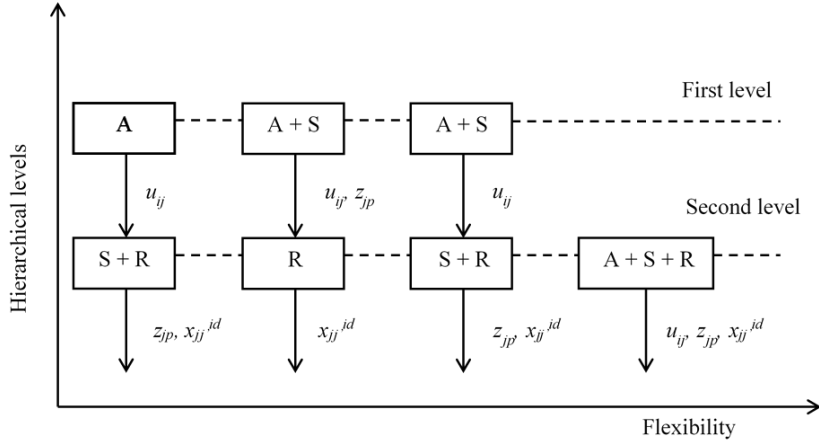
11

Figure 1: Rigid models vs flexible models

pattern-based two-phase method is equity driven while the second phase can be either cost or again equity driven. The equity refers to a fair distribution of the workload among operators and the optimization function is thus a balancing criterion. At that point, two equally prominent points of view can be considered: we can focus either on the most busy operator and minimize his/her workload, or alternatively we can focus on the least busy operator and maximize his/her workload. In the following, the character "W" is used to denote the workload while indexes "MB" and "LB" stand respectively for *most busy* and *least busy*. Consequently, W_{MB} and W_{LB} are used respectively to denote the two above mentioned balancing criteria. In contrast, when the focus is on cost control the optimization will be guided by the minimization of the overall distance traveled by all operators. The letter "C" is used to denote the cost.

For each of the four methods represented in Figure 1 we use the notation $v_1(f_1)|v_2(f_2)$ to describe the two optimization problems addressed respectively in the two phases, where v_i for $i = 1, 2$ refers to the decision variables involved (see column "Label" in Table 1) while f_i for $i = 1, 2$ refers to the criterion used to guide the corresponding optimization problem. As an example, $A(W_{MB})|S+R(C)$ refers to a two-phase method where the first phase solves an assignment problem and is guided by a balancing criterion focusing on the most busy operator, whereas the second phase solves the scheduling and routing problems by using cost minimization as a guiding criterion.

We propose four variants of $A|S+R$, namely $A(W_{MB})|S+R(W_{MB})$, $A(W_{MB})|S+R(C)$, $A(W_{LB})|S+R(W_{LB})$, $A(W_{LB})|S+R(C)$, where the first phase is guided by two balancing criteria, while the second phase can be guided by a cost criterion or by a balancing criterion in accordance with the one chosen at the upper level. The

12

same considerations hold for $A + S|R$ and $A + S|S + R$. Finally, we consider three variants of the the single-phase method $.|A + S + R$, which can be guided by one of the two balancing criteria or by the minimization of cost.

Summarizing, we consider four alternative pattern-based methods for a total of 15 variants. For more clarity, the used notation is summarized in Table 2.

Table 2: Summary of used notation

Set	Description
$\mathscr{H}$	planning horizon
$\mathscr{N}$	set of patients
$\mathscr{A}$	set of possible trajectories
$\mathscr{K}$	set of skill levels
$\mathscr{P}$	set of patterns
$\mathscr{O}$	set of available operators
$\mathscr{O}_d$	subset of operators available on day d
$\mathscr{S}_k$	subset of operators having at least skill k
Parameter	**Description**
a_i	capacity of operator i
r_j	care plan of patient j
r_{jk}	number of visits of skill k required by patient j
$t_{jj'}$	distance between patients j and j'
t'_j	care service time at patient j
τ_j	average traveling time to reach patient j
α_j	total volume of care required by patient j

4 Pattern-based methods

The aim of this section is to give evidence of the structure of the problem addressed and to show how the four methods presented, and their variants, capture such a structure with different levels of flexibility. In addition, we list all those issues that are common to the methods and we emphasize the differences between them.

As anticipated in the previous section, the main decisions to take are assignment, scheduling and routing decisions which, in all of the four methods, are modeled respectively by means of the three following sets of variables:

$$u_{ij} = \begin{cases} 1 \text{ if operator } i \text{ is assigned to patient } j \\ 0 \text{ otherwise} \end{cases} \quad i \in \mathscr{O}, j \in \mathscr{N}$$

13

$$z_{jp} = \begin{cases} 1 \text{ if pattern } p \text{ is assigned to patient } j \\ 0 \text{ otherwise} \end{cases} \quad j \in \mathscr{N}, p \in \mathscr{P}$$

$$x_{jj'}^{id} = \begin{cases} 1 \text{ if operator } i \text{ travels along } (j, j') \text{ on day } d \\ 0 \text{ otherwise} \end{cases} \quad (j, j') \in \mathscr{A}, d \in \mathscr{H}, i \in \mathscr{O}_d.$$

Analogously, the set of constraints that describe HRPHHC can be grouped in three blocks: assignment, scheduling and routing constraints. Specifically, the constraints regarding assignment decisions have to guarantee that: (*i*) exactly one operator of adequate skill is assigned to each patient; (*ii*) the workload of the operators is controlled. The constraints regarding the scheduling decisions have to guarantee that: (*i*) for each patient, each visit envisaged in her/his care plan is scheduled in the planning horizon. Finally, the constraints regarding routing decisions have to guarantee that: (*i*) each operator on each scheduled day of the planning horizon performs a route that starts from the depot and ends to the depot, i.e. subtours disconnected from the depot are not allowed; (*ii*) the daily workload of each operator, expressed both in terms of visit time and travel time, is controlled.

Clearly all the decisions and constraints have to be coordinated. At first, we have to assure that if a visit of a given patient j is allocated on day d one operator of adequate skill and available on day d performs the visit, i.e. the route of the operator on that day has to include patient j. These are linking constraints that bind together scheduling and routing decisions. In addition, we have to assure that an operator can visit a patient only if she/he has been assigned to that patient (linking between routing and assignment variables). Figure 2 shows the three different types of decisions with their constraints as well as the linking constraints between assignment and scheduling and between scheduling and routing.

The way the decisions are coordinated characterizes the four methods presented. The most flexible method consists of a mathematical formulation where the set of constraints comprises the above mentioned three blocks of constraints in addition to the linking constraints. When the linking constraints are relaxed by elimination, a decomposition method arises which is organized in two phases according to the scheme shown in the previous section. In this sense, the models associated with the decomposition methods can be viewed as relaxations of the single-phase model. Indeed, the more flexible the more information-aware the method is. For this reason, we start presenting the single-phase method $.|A + S + R$; then we will show the decomposed approaches discussing the mathematical formulation used in each phase and highlight the relationships with respect to the single-phase method.

The objective function used, either a cost criterion or an equity criterion, is common to all the methods and is described in details for the most flexible one.

14

Its application to the decomposition methods is straightforward and potential peculiarities are discussed when required.

Likewise the objective function used, another key issue that is common across all the methods proposed is the way skills are managed, either independently or hierarchically.
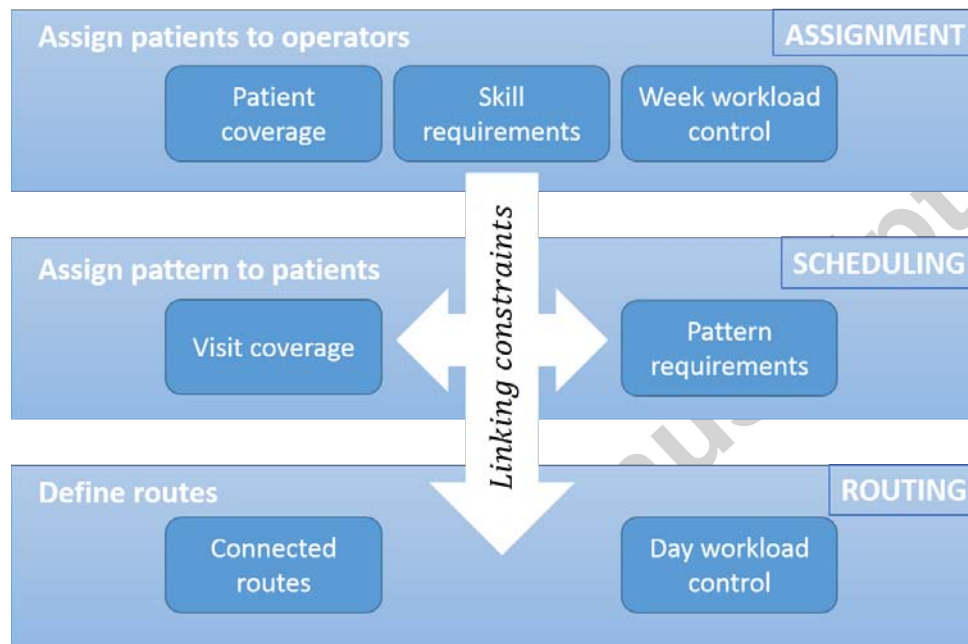

Figure 2: Schema of subproblems: assignment, scheduling and routing

4.1 The most flexible method: $.|A + S + R$

In the most flexible method the model is described by the following feasibility set, which is shortly described next. For details we refer to Cappanera and Scutellà

(2015).

$$\sum_{i \in \mathcal{O}} u_{ij} = 1 \qquad\qquad \forall j \in \mathcal{N} \qquad\qquad (1)$$

$$\sum_{p \in P} z_{jp} = 1 \qquad\qquad \forall j \in \mathcal{N} \qquad\qquad (2)$$

$$\sum_{j \in \mathcal{N} \cup \{0\}} x_{jj'}^{id} = \sum_{j \in \mathcal{N} \cup \{0\}} x_{j'j}^{id} \qquad\qquad \forall j' \in \mathcal{N}, \forall d \in \mathcal{H}, \forall i \in \mathcal{O}_d \qquad (3)$$

$$D_{id} = \sum_{(j,j') \in \mathcal{A}} (t_j' + t_{jj'}) \cdot x_{jj'}^{id} \le a_i \qquad\qquad \forall d \in \mathcal{H}, \forall i \in \mathcal{O}_d \qquad (4)$$

$$\sum_{j \in \mathcal{N}} y_{0j}^d = \sum_{j \in \mathcal{N}} \sum_{p:p(d) \ge 1} z_{jp} \qquad\qquad \forall d \in \mathcal{H} \qquad\qquad (5)$$

$$\sum_{j \in \mathcal{N} \cup \{0\}} y_{jj'}^d - \sum_{j \in \mathcal{N} \cup \{0\}} y_{j'j}^d = \sum_{p:p(d) \ge 1} z_{j'p} \quad \forall j' \in \mathcal{N}, \forall d \in \mathcal{H} \qquad (6)$$

$$\sum_{j \in \mathcal{N} \cup \{0\}} \sum_{i \in \mathcal{O}} x_{jj'}^{id} \le \sum_{p:p(d) \ge 1} z_{j'p} \qquad\qquad \forall j' \in \mathcal{N}, \forall d \in \mathcal{H} \qquad (7)$$

$$\sum_{j \in \mathcal{N} \cup \{0\}} \sum_{i \in \mathcal{S}_k} x_{jj'}^{id} \ge \sum_{p:p(d) = k} z_{j'p} \qquad\qquad \forall j' \in \mathcal{N}, \forall d \in \mathcal{H}, \forall k \in \mathcal{K} \qquad (8)$$

$$x_{jj'}^{id} \le u_{ij} \qquad\qquad \forall (j, j') \in \mathcal{A}, \forall d \in \mathcal{H}, \forall i \in \mathcal{O}_d \quad (9)$$

$$u_{ij} \le \sum_{j \in \mathcal{N} \cup \{0\}} \sum_{d \in \mathcal{H}} x_{jj'}^{id} \qquad\qquad \forall j' \in \mathcal{N}, \forall i \in \mathcal{O} \qquad (10)$$

$$y_{jj'}^d \le n \sum_{i \in \mathcal{O}_d} x_{jj'}^{id} \qquad\qquad \forall (j, j') \in \mathcal{A}, \forall d \in \mathcal{H} \qquad (11)$$

Constraints (1) are the operator-patient assignment constraints and they assure that exactly one operator is assigned to each patient during the planning horizon. This is included in order to guarantee continuity of the care. Constraints (2) are the scheduling constraints and they assure that each patient is assigned exactly to a pattern. Then the routing constraints follow. Specifically, constraints (3) are the classical flow conservation constraints on the routing variables. Constraints (4) assure that the workload of each operator in each day, expressed as the sum of the service times (i.e., t_j' is the care service time at node j) and the traveling times, does not exceed the duration of a workday, i.e. a_i. For each operator i, the deterministic capacity a_i corresponds to the maximum amount of time (on a single day) that the operator can work according to his/her working contract. Constraints (5) and (6) are the flow conservation constraints on the auxiliary y variables, which are introduced to avoid subtours in the model solutions (Cappanera et al. (2013)). Specifically, each $y_{jj'}^d$ is an auxiliary flow variable that represents the number of patients visited after patient j by the operator moving along (j, j') on day d. Constraints (5) and (6) also guarantee the correct linking between scheduling decisions

16

and auxiliary flow variables. Constraints (7) and (8) link together scheduling and routing variables. Specifically, constraints (7) state that (exactly) one operator per day can visit patient j only if a visit has been scheduled on that day for node j. Constraints (8) guarantee that, on day d, exactly one operator, of adequate skill, must visit patient j if a service of skill k has been scheduled for j on day d, for each k. Due to the hierarchical skill management, a visit requiring skill k may be operated by each of the operator owing a skill at least k; such a set is denoted by $\mathscr{S}_k$. In particular, the least skilled operators can perform only visits of skill 1 (case $k = 1$), whereas the most skilled operators can perform all types of visits (case $k = K$). Constraints (9) and (10) link together assignment and routing variables. Specifically, constraints (9) guarantee that an operator can visit a patient only if he has been assigned to that patient, while constraints (10) force variables u_{ij} to zero if operator i never visits patient j during the planning horizon. Finally, constraints (11) link together routing variables and auxiliary flow variables.

Further observe that, for each $j \in \mathscr{N}$, a pattern variable z_{jp} can assume a value other than zero only if:

$$|\{d : p(d) = k\}| = r_{jk} \quad \forall k \in \mathscr{K}. \tag{12}$$

Therefore, in the preprocessing phase $z_{jp} = 0$ if any of the K constraints (12) is not satisfied.

The way the skills are managed is reflected in the set of operators eligible for a given patient j. In constraints (8) skills are managed jointly, and therefore the patients having $k \in \mathscr{K}$ as the highest level of skill appearing in their care plan have to be assigned to an operator with skill greater than or equal to k, i.e. belonging to the set $\mathscr{S}_k$. On the other hand, when the skills are managed independently, the patients having $k \in \mathscr{K}$ as the highest level of skill appearing in their care plan have to be assigned to an operator with skill exactly k. In such a case, the linking constraints (8) have to be modified accordingly. Thus, properly defining the set of operators eligible for a given patient, we can handle both types of skill management schemes by method $.|A + S + R$.

As discussed in Section 3.2, the objective function represents an equity criterion or a cost criterion. When the focus is on equity, operators' utilization rates play a crucial role, where the *utilization rate* of an operator is the ratio between the actual workload of the operator and his/her capacity. The operator workload is composed of the total service times plus the traveling times to reach patient homes. If the point of view of the most busy operator is assumed (W_{MB}), the aim is to minimize the maximum operator utilization rate and the objective function is formulated as follows:

$$W_{MB} \quad \min \quad m_u$$

$$\frac{\sum_{d \in \mathcal{H}} D_{id}}{H \cdot a_i} \leq m_u \qquad \qquad \forall i \in \mathcal{O} \qquad (13)$$

where D_{id} are defined in constraints (4), while m_u is an upper bound to the maximum utilization rate of the operators. Indeed, by minimizing m_u we minimize the maximum utilization rate of the operators.

Symmetrically, when the focus is on the least busy operator (W_{LB}) the maximization of the minimum operator utilization rate is formulated as follows:

$$W_{LB} \quad \max \quad m_l$$

$$\frac{\sum_{d \in \mathcal{H}} D_{id}}{H \cdot a_i} \geq m_l \qquad \qquad \forall i \in \mathcal{O} \qquad (14)$$

where m_l expresses a lower bound to the minimum utilization rate of the operators to be maximized.

Finally, when the focus is on cost control the following objective function is used:

$$C \quad \min \quad \sum_{i \in \mathcal{O}} \sum_{d \in \mathcal{H}} \sum_{(j,j') \in \mathcal{A}} t_{jj'} x_{jj'}^{id}.$$

4.2   The most rigid method: $A|S+R$

This section aims at presenting mathematical models for the two-phase method $A|S+R$.

First phase. As emphasized earlier, the assignment process (first phase of the method) consists in assigning operators to patients in such a way that skill constraints are satisfied and the workload of the operators is controlled. Since scheduling and routing aspects are disregarded in this phase, the requirement α_j of each patient j is calculated as the total amount of care volume required by j, for all levels of skill, over the planning horizon and by an estimate of the travelling times. Specifically,

$$\alpha_j = \sum_{k=1}^{K} r_{jk}(t'_j + \tau_j), \forall j \in \mathcal{N}, \qquad (15)$$

where t'_j, previously introduced, is the care service time at node j, whereas τ_j denotes the average traveling time to reach patient j from each node other than j, included the depot:

18

$$\tau_j = \frac{1}{n} \sum_{j \neq j'} t_{jj'}. \tag{16}$$

From the definitions above, α_j estimates the time that the operator assigned to j will dedicate to that patient during the planning horizon.

Assignment variables and assignment constraints are exactly the same as the corresponding ones in the single-phase method. Also in this case, when the skills are managed independently, the set of operators eligible for a given patient j is given by all those operators with a skill equal to the maximum skill required in the care plan of j; when skills are managed hierarchically, an operator is eligible when she/he has a skill greater than or equal to the maximum skill required by the patient.

As discussed before, the first phase is equity driven in the whole range of methods. Differently from what happens in the single-phase method, here the workload of operator i is expressed by means of the estimated demand, i.e.

$$\frac{\sum_{j \in \mathcal{N}} \alpha_j u_{ij}}{H \cdot a_i}, \forall i \in \mathcal{O}. \tag{17}$$

When skills are managed independently, the assignment problem decomposes by skill and a model is solved for each skill level, in a separate way. Under the other variant, the assignment decisions are taken simultaneously for all patients and operators. In other words, a single assignment problem is solved involving all operators to handle all patient requirements.

Since scheduling decisions are not considered at this stage, the assignment models in this method are single period and they can be viewed as special cases of those presented in (Lanzarone et al., 2012).

Second phase. In the second phase of the $A|S + R$ method, the aim is to schedule the patient requests during the planning horizon and at the same time to determine the daily routes of each operator on each day of the horizon.

It is important to point out that the output of the assignment problem solved in the first phase, i.e. the set of patients assigned to each operator, becomes the input of the second phase. Assignment decisions do not come into play anymore; consequently, the model that characterizes the second phase decomposes operator wise.

The constraints of the model in the second phase are a relaxation of the constraints describing the single-phase model presented in Section 4.1. Specifically, the set of constraints is given by scheduling, routing and related linking constraints, i.e. by constraints (2), (3), (4), (5), (6), (7), (8). Also constraints (9) and (10) are present: in each of them however the value of variables u_{ij} is fixed

19

according to the solution of the first phase. With regards to the linking constraints between flow and routing variables, we can refine constraints (11) in the single-phase model as follows. Let $\mathcal{N}_i$ be the subset of the patients who have been assigned to operator i as a result of the assignment phase. In addition observe that, for each operator i, the routing decisions are taken in the subgraph of G induced by the node set $\mathcal{N}_i$. Consequently, the linking constraints (11) in the single-phase method can be replaced by the following constraints:

$$y_{jj'}^d \leq |\mathcal{N}_i| x_{jj'}^{id} \qquad \forall (j, j') \in \mathscr{A}, \forall d \in \mathscr{H} : i \in \mathscr{O}_d. \qquad (18)$$

Thus, the model in the second phase derives straightforwardly by the single-phase model; it is only necessary to give some details on how certain sets and parameters can be refined. In particular we can associate a set of patterns, say $\mathscr{P}_i \subseteq \mathscr{P}$, with each operator i, by excluding from $\mathscr{P}$ all those patterns that contemplate an activity on a day in which operator i is not available. Specifically, for each operator i, $\mathscr{P}_i$ must be selected in such a way that $p \in \mathscr{P}_i$ implies $p(d) = 0$ if $i \notin \mathscr{O}_d$. The scheduling constraints (2) can thus be replaced by the following constraints:

$$\sum_{p \in \mathscr{P}_i} z_{jp} = 1 \qquad \forall j \in \mathcal{N}_i. \qquad (19)$$

The objective function used in the second phase can be alternatively either a cost criterion or an equity criterion as discussed in Section 3.2. When an equity criterion is selected both in the first and in the second phase accordingly, the utilization rate of the operators is the one returned by the second phase of the decomposition method.

4.3   The intermediate methods: $A + S|R$ and $A + S|S + R$

First phase. Both of the intermediate methods $A + S|R$ and $A + S|S + R$ are characterized by a first phase where the assignment of operators to patients is made on a richer model with respect to the rigid method $A|S + R$. In the rigid method the assignment is made by looking at the care plan of a patient as a whole, i.e. guaranteeing that the total time spent by an operator in the visits during the planning horizon does not exceed the weekly availability of the operator. This type of patient-operator assignment can very likely generate infeasibilities when the daily workload of an operator is considered. For these reasons, the intermediate methods incorporate scheduling decisions in the first phase and consider assignment variables u_{ij}^d that are disaggregated by day with respect to the variables u_{ij}. More formally,

$$u_{ij}^d = \begin{cases} 1 \text{ if operator } i \text{ visits patient } j \text{ on day } d \\ 0 \text{ otherwise} \end{cases} \quad j \in \mathcal{N}, d \in \mathcal{H}, i \in \mathcal{O}_d.$$

The mathematical formulation of the problem solved in the first phase is the following:

$$\sum_{i \in \mathcal{O}} u_{ij} = 1 \qquad \forall j \in \mathcal{N} \qquad (20)$$

$$\sum_{p \in P} z_{jp} = 1 \qquad \forall j \in \mathcal{N} \qquad (21)$$

$$D_{id} = \sum_{j \in \mathcal{N}} (t'_j + \tau_j) \cdot u_{ij}^d \leq a_i \qquad \forall d \in \mathcal{H}, \forall i \in \mathcal{O}_d \qquad (22)$$

$$\sum_{i \in \mathcal{O}} u_{ij}^d \leq \sum_{p:p(d) \geq 1} z_{jp} \qquad \forall j \in \mathcal{N}, \forall d \in \mathcal{H} \qquad (23)$$

$$\sum_{i \in \mathcal{S}_k} u_{ij}^d \geq \sum_{p:p(d)=k} z_{jp} \qquad \forall j \in \mathcal{N}, \forall d \in \mathcal{H}, \forall k \in \mathcal{K} \qquad (24)$$

$$u_{ij}^d \leq u_{ij} \qquad \forall j \in \mathcal{N}, \forall d \in \mathcal{H}, \forall i \in \mathcal{O}_d \qquad (25)$$

Constraints (20) are the assignment constraints and correspond to constraints (1) in the single-phase model. Constraints (21) are the scheduling constraints and they are the same as constraints (2). Constraints (22) control the daily workload of the operators and make use of τ_j to estimate the travel time to reach patient j, as it happens in the first phase of the method $A|S+R$. Constraints (23), (24) and (25) link together assignment and scheduling decisions; specifically, constraints (23) state that (exactly) one operator per day can visit patient j only if a visit has been scheduled on that day for node j. Constraints (24) guarantee that, at each day, exactly one operator of adequate skill must visit a patient for whom a visit is scheduled on that day, while constraints (25) guarantee that an operator can visit a patient only if she/he has been assigned to that patient.

As in the previous methods, the skills can be managed either jointly, as realized by constraints (24), or independently. When independent skills are assumed the problem decomposes by skill. Also for the intermediate methods the first phase is equity driven.

Second phase. The two intermediate methods share the first phase but they differ in regards to the second phase; specifically, in the second phase of the method $A+S|S+R$, which is the more flexible of the two, scheduling and routing decisions are taken exactly as it happens in the second phase of the method $A|S+R$. Only assignment decisions taken in the first phase are maintained while scheduling decisions come again into play. On the contrary, in the method $A+S|R$ assign-

21

ment and scheduling decisions taken in the first phase are fixed and the second phase consists only in a routing phase.

Observing the structure of the single-phase model we have that, by fixing the assignment variables u_{ij} and the scheduling variables z_{jp}, the routing problem consists in determining a route for each operator i and for each day d on a reduced graph which is made by all the nodes corresponding to patients assigned to i on day d. In addition, the daily workload of the operator must not exceed her/his daily capacity. We observe that the routing problem may further decompose according to the objective function used. Specifically, when a cost criterion is used, the routing problem decomposes by day and by operator thus resulting in a Capacitated Traveling Salesman Problem for each operator and for each day; in contrast, when a balancing objective function is used the routing problem cannot be decomposed. However, whatever the objective function is, the routing problem can be solved by means of the mathematical formulation of the single-phase model (method $.|A + S + R$) where the assignment and the scheduling variables are fixed according to the solution obtained in the first phase.

5   Computational results

In this section some computational results are presented on a set of instances generated from real data. They are analyzed with respect to the used methods (i.e. single-phase and two-phase methods) under different skill management strategies (i.e. hierarchical versus independent skill management) and optimization criteria (i.e. the cost criterion C, and the balancing criteria W_{LB} and W_{MB}). With respect to the description provided in Section 4, at the implementation level all the methods have been enhanced with the symmetry management valid inequalities presented in Cappanera and Scutellà (2015). Furthermore, with the exception of the variants using the cost criterion C, also the exploit-cluster valid inequalities have been added to the models since, as shown in Cappanera and Scutellà (2015), they represent an effective tool to take the travel time under control.

Main conclusions are derived based on the computational efficiency (in terms of the optimality gap and of the computational time) and the quality of some solution indicators (i.e. the operator utilization factor and the travel time).

The goal of the numerical experimentation is to assess the advantages of the proposed pattern-based decomposition methods wih respect to the single-phase method introduced in Cappanera and Scutellà (2015). In fact, such a method is very powerful and flexible, but it may be not suitable for instances of large size. We will show that some of the proposed decomposition methods allow one to improve the computational efficiency of the single-phase approach, by returning very good quality solutions (in terms of the optimality gap and of the stated solution

indicators) in a consistently faster computational time, also for instances of large size.

The main achievement of the computational experience presented in this section is therefore that the proposed pattern-based decomposition is a valuable approach to improve on the computational efficiency of the single-phase method, without substantially deteriorating its flexibility, i.e. the quality of the returned solutions. Based on this promising result, more sophisticated exact and heuristic algorithms, relying on the proposed pattern-decomposition, will be investigated in order to be able to face with even larger Home Care instances.

In the following subsections, firstly the real data and the generated instances will be described. Then, an analysis of the results related to the computational efficiency will be presented. Finally, the quality of the computed solutions will be discussed.

5.1   Real data and generated instances

The real data have been obtained from one of the largest Italian public HHC providers that operates in Lombardia region and covers approximately 800 km^2 with three independent divisions. The analysis in this work refers to the largest division and its 7 municipalities. In particular, this provider serves mainly palliative and non-palliative (i.e. ordinary) patients. Each generated instance has the following characteristics: the set of the municipalities (i.e., the geographical region of the service), the time horizon (i.e., 6 days), the set of the patients, the set of the operators, the categories of the patient requests and the operator skills (i.e., *ordinary* patient (skill 1) or *palliative* patient (skill 2)), the operators' capacity and the set of the patterns. Specifically, for each instance these characteristics are gathered, in regards to the patients, by selecting arbitrarily a week in 2004 or 2005 from the historical database of the provider. Then, as the next step, the demands of the patients have been randomly generated by using the method proposed in (Lanzarone et al., 2010) for the selected week.

A total of 34 instances have been generated, which are grouped in five sets. The first set consists of 10 small-size instances with 60 patients and 4 operators. The second set comprises 10 medium-size instances with 100 patients and 7 operators. The third group includes 10 large-size instances consisting of 200 patients and 10 operators. Finally, the pool of instances also comprises 2 big instances with 250 patients and 2 big instances with 300 patients, with 13 and 16 operators, respectively. In each instance, both skill 1 and skill 2 operators are included with workday durations (i.e. the parameters a_i) of five to eight hours. The traveling times (i.e. the parameters $t_{jj'}$) between patients located in different municipalities have been calculated via Google maps, whereas the traveling time between

patients located in the same municipality has been assumed to be equal to three minutes. Both skills require clinical competencies, thus the same service time (i.e. the parameters t'_j) has been assumed and set equal to 45 minutes for all patients. By considering the patterns, they have been generated with the use of the flow-based policy proposed by Cappanera and Scutellà (2013). The instance features are provided in Table 3.

Table 3: Instance features

Instance	Number of Skill1 Patients	Number of Skill2 Patients	Number of Skill1 Visits	Number of Skill2 Visits	Number of Patterns	Number of Skill1 Operators	Number of Skill2 Operators
60-1	51	9	69	28	9	2	2
60-2	51	9	76	28	8	2	2
60-3	48	12	67	38	8	2	2
60-4	50	10	70	28	8	2	2
60-5	44	16	69	46	7	2	2
60-6	53	7	72	23	10	2	2
60-7	52	8	74	21	9	2	2
60-8	52	8	75	26	8	2	2
60-9	56	4	80	7	7	2	2
60-10	49	11	62	36	7	2	2
100-1	84	16	115	46	8	4	3
100-2	82	18	122	48	9	4	3
100-3	86	14	124	39	9	4	3
100-4	88	12	123	39	10	4	3
100-5	80	20	111	66	10	4	3
100-6	94	6	172	21	8	4	3
100-7	82	18	126	51	10	4	3
100-8	92	8	127	23	8	4	3
100-9	93	7	128	19	8	4	3
100-10	89	11	113	38	8	4	3
200-1	168	32	199	104	16	6	4
200-2	165	35	253	96	16	6	4
200-3	171	29	246	92	18	6	4
200-4	174	26	253	71	12	6	4
200-5	159	41	226	132	16	6	4
200-6	174	26	241	83	12	6	4
200-7	180	20	256	66	14	6	4
200-8	177	23	258	77	13	6	4
200-9	178	22	273	64	11	6	4
200-10	178	22	259	58	10	6	4
250-1	210	40	292	131	14	8	5
250-2	213	37	315	137	14	8	5
300-1	250	50	405	152	13	10	6
300-2	258	42	362	152	12	10	6

The experiments on the medium, large and big instances were carried out on an Intel(R) Core(TM) i7-4770 (CPU 3.40 GHz), and CPLEX 12.6.1 was used as the solver. The experiments on the small-size instances, instead, were carried out on an AMD Opteron(TM) Dual Core Processor 246 (CPU 1.9 GHz), by using a less recent version of CPLEX.

In all the cases a time limit of 12 hours and a memory limit of 1 GB have been

imposed. Some additional experiments have been performed with time limits of 15 minutes, 1 hour and 6 hours, respectively, in order to emphasize the capability of some decomposition approaches to efficiently compute, in a reduced amount of time, solutions near to the optimal ones. Hereafter the computational times are always expressed in seconds of CPU time. Furthermore, when memory limit is exceeded, this is identified with a $\star$.

5.2   Efficiency analysis

In this section we analyze the computational efficiency of the pattern-based decomposition by solving the 34 instances with different optimization criteria and skill management policies. The considered indicators are the *computational time* and the *% optimality gap* (i.e. the percentage relative error of the best solution value with respect to the best bound computed by the solver).

First of all, it is important to emphasize that, for the considered pool of instances, no feasible solutions were determined under the independent skill management policy. On the contrary, the hierarchical skill management policy was able to provide feasible solutions for most of the cases. Thus, the hierarchical skill management policy appears to be relevant to deal with real situations of the HHC services. All the results in this study therefore refer to this policy.

Furthermore, as far as the most rigid decomposition approach is concerned, i.e. $A|S+R$, no results will be provided since this method failed to provide feasible solutions whatever optimization criterion was used in both phases.

Let us first analyze the experimental results on the small-scale instances. As expected, the computational time of the single-phase method is much larger than that of the two-phase methods (see Figure 3). When the cost is minimized, the single-phase method almost always reached the time limit. However, by considering the balancing optimization criteria, the boxplot shows that the single-phase method performed with lower computational times, especially when the W_{LB} criterion is adopted.

The gaps reported in Figure 4 confirm the difficulties of the single-phase method when a cost criterion is adopted. In this case the average gap, which is 50.5%, is not acceptable. On the contrary, the gaps reduce remarkably when a balancing criterion is used, and they are smaller with W_{LB} than with W_{MB}, according to the computational results in Cappanera and Scutellà (2015). Specifically, under the W_{LB} optimization criterion the optimality gap is below 6.30% for all the small-size instances but one, and it is below 26.65% under W_{MB} showing, however, two cases of infeasibility.

The proposed pattern-based decomposition methods proved to be much more efficient, as shown by Figures 3 and 5. Figure 6 plots the related % optimality
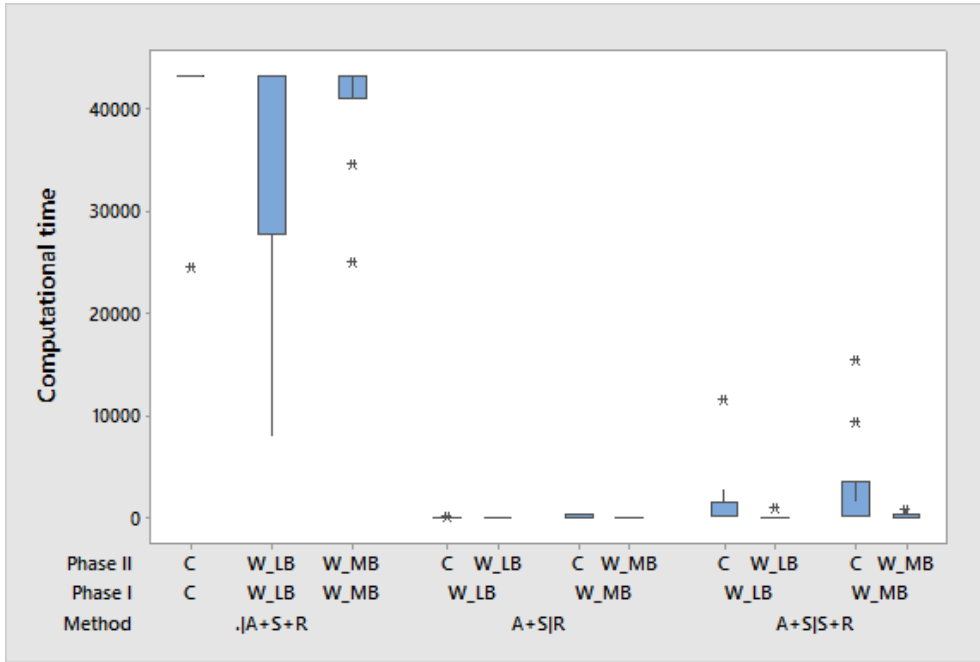
Figure 3: Boxplot of computational times (in seconds) for single-phase and two-phase methods (small-scale instances).

gaps. The following overall considerations can be done: (i) the method $A+S|S+R$ has a larger computational time compared to $A+S|R$, but it allows to obtain better optimality gaps, (ii) minimizing costs in the second phase takes largely more time than balancing, (iii) balancing the utilization factors by maximizing the least busy operator (W_{LB}) in the first phase requires a lower computational time compared to minimizing the most busy operator (W_{MB}).

In particular, the pattern-based decomposition method $A+S|S+R$ showed to be advantageous both with respect to the single-phase approach, and also with respect to the alternative pattern-based decompositions. In fact, it guarantees better optimality gaps under all the combinations of the optimization objectives in the two phases of the approach, with a just slightly increase in the computational time. The method is however very fast in most cases. Therefore, the first achievement is that the tested small-size instances are efficiently affordable by means of $A+S|S+R$, that is when patterns are used in both phases of the decomposition approach.

More in detail, by minimizing the cost in the second phase the computational time of $A+S|S+R$ is below 116.10 seconds for the majority of the cases (by considering both kinds of balancing criteria in the first phase), and the optimality gap is almost always below 0.84%. By considering the W_{LB} criterion in both
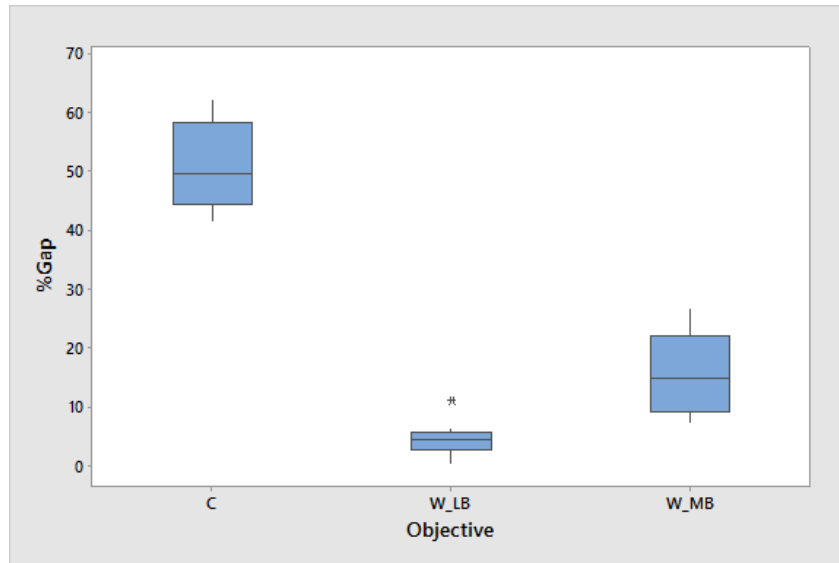
26

Figure 4: Boxplot of the % Gap of single-phase method (small-scale instances).

phases, the percentage gap is zero for all the instances but one, in less than 3.51 seconds, whereas by selecting the W_{MB} criterion in both phases, the optimality gap is always less than 0.81%, in less than 22.58 seconds for all instances but four (with a maximum of 818.23 seconds).

According to these promising results, $A + S|S + R$ has been reputed the best decomposition methods for solving the even larger instances. Therefore, the efficiency and the quality results reported for the instances with 100, 200, 250 and 300 patients refer to such a method.
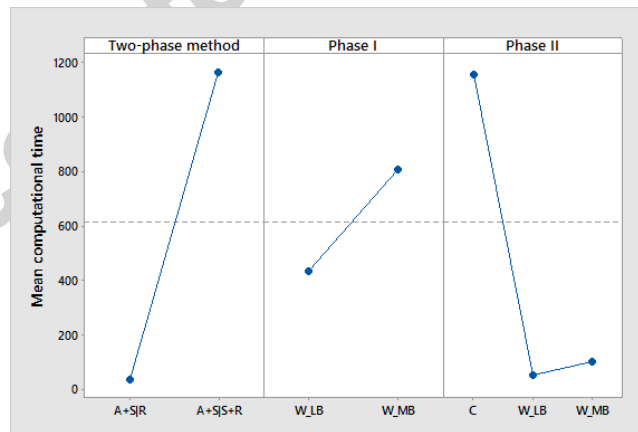
Figure 5: Efficiency analysis of two-phase methods: computational time in seconds (small-scale instances)
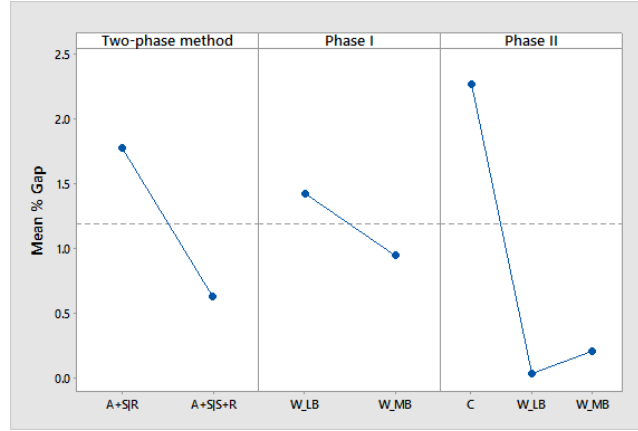
Figure 6: Efficiency analysis of two-phase methods: gap (small-scale instances)

Consider now the larger instances. As expected, with the medium-scale (i.e. 100 patients), the large-scale (i.e. 200 patients) and the big (i.e. 250 and 300 patients) instances, the single-phase method failed to provide feasible solutions almost everywhere, within the considered time and memory limit.

On the other hand, the pattern-based decomposition method $A+S|S+R$ proved to be capable to address also these larger instances in an efficient way. Specifically, by considering the medium-scale instances: (i) using the W_{LB} criterion in both phases, the decomposition approach was able to compute an optimal solution in all cases but one (with gap 0.12%) in less than about 40 seconds (see Table 4) (ii) using the W_{MB} criterion in both phases, the gap was less than 0.40% in most cases, with one infeasibility, in a greater computational time with respect to W_{LB} (see Figures 7 and 8, showing the bloxpot time and gap, respectively) (iii) using the cost criterion in the second phase increased the computational time substantially, according to the results obtained on the small-scale instances; furthermore, also the percentage gaps augmented (see Figures 7 and 8).

The efficient behavior of $A+S|S+R$ has been confirmed on the large-scale instances: (i) using the W_{LB} criterion in both phases, the decomposition approach was able to compute an optimal solution in almost all the cases, with just one infeasibility, most often in less than 4 minutes (see Table 5) (ii) the W_{MB} criterion, used in both phases, showed to be less suitable for these larger instances, returning very small gaps in some cases, but 5 infeasibilities; also for these instances, the computational time was greater than the one of W_{LB} (see Figures 9 and 10, showing the bloxpot time and gap, respectively) (iii) using the cost criterion in the second phase increased both the computational time and the gap; anyway, feasible solutions were always determined, in the majority of the cases with a gap less than 0.20% (see Figure 10).

28

Table 4: A+S(W_{LB})|S+R(W_{LB}): behavior on the medium-scale instances (time in seconds)

Instance	LPTime	LPValue (root)	IPValue	%Gap	IPTime
100-1	0.34	0.5776	0.5776	0.00	6.38
100-2	0.66	0.6011	0.5911	0.00	8.20
100-3	0.74	0.6031	0.6004	0.12	9.95
100-4	0.67	0.5678	0.5678	0.00	35.05
100-5	0.20	0.5650	0.5650	0.00	8.12
100-6	1.68	0.7569	0.7561	0.00	32.45
100-7	0.84	0.5967	0.5967	0.00	6.05
100-8	1.02	0.5761	0.5761	0.00	40.43
100-9	0.36	0.5524	0.5524	0.00	13.65
100-10	0.91	0.5406	0.5406	0.00	10.70

Finally consider the big instances. Using the W_{LB} criterion in both phases, $A + S|S + R$ computed an optimal solution in two cases, without determining any feasible solution in the remaining two cases (see Table 6). Again, the W_{MB} criterion, used in both phases, showed to be less suitable for addressing the larger instances, with 3 infeasibilities over 4. Using the cost criterion in the second phase increased both the computational time and the gap; however, a feasible solution was computed in all cases (see Figures 11 and 12, showing the bloxpot time and gap, respectively).

The main achievement is therefore that, as for the small-case instances, the version of $A + S|S + R$ using the balancing criterion W_{LB} in both phases is very efficient in addressing the medium, the large and the big instances, returning, in most cases, solutions with a very low percentage gap in a short or however reasonable time. The versions using the cost criterion in the second phase, although required an increased computational time and, often, returned solutions with larger gaps, proved to be very robust, being capable to compute a feasible solution in all the tested cases.

Some additional experiments have been performed with time limits of 15 minutes, 1 hour and 6 hours, respectively, in order to emphasize the capability of $A + S|S + R$ to efficiently compute, in a reduced amount of time, solutions near to the optimal ones. The results are very interesting. In fact, on 76 tests for which a feasible solution was computed, within 15 minutes $A + S|S + R$ computed 43 solutions with a percentage gap less than 5%. This number increases to 47, so for

Table 5: A+S(W_{LB})|S+R(W_{LB}): behavior on the large-scale instances (time in seconds)

Instance	LPTime	LPValue (root)	IPValue	%Gap	IPTime
200-1	11.95	0.6990	0.6990	0.00	2138.71
200-2	6.59	0.8173	0.8173	0.00	187.87
200-3 $\star$	19.03	0.7983	n.a.	n.a.	3358.60
200-4 $\star$	7.06	0.7549	0.7533	0.12	31367.41
200-6	6.43	0.7621	0.7621	0.00	384.79
200-7	5.35	0.7580	0.7580	0.00	496.93
200-8$\star$	9.93	0.7885	0.7871	0.15	12535.61
200-9	1.51	0.7944	0.7933	0.03	48.32
200-10	7.26	0.7646	0.7646	0.00	402.66

Table 6: A+S(W_{LB})|S+R(W_{LB}): behavior on the big instances (time in seconds)

Instance	LPTime	LPValue (root)	IPValue	%Gap	IPTime
250-1 $\star$	14.94	0.7517	n.a.	n.a.	4265.11
250-2	7.36	0.7956	0.7956	0.00	11257.93
300-1 $\star$	9.04	0.8157	n.a.	n.a.	3237.59
300-2	4.66	0.7290	0.7290	0.00	141.32
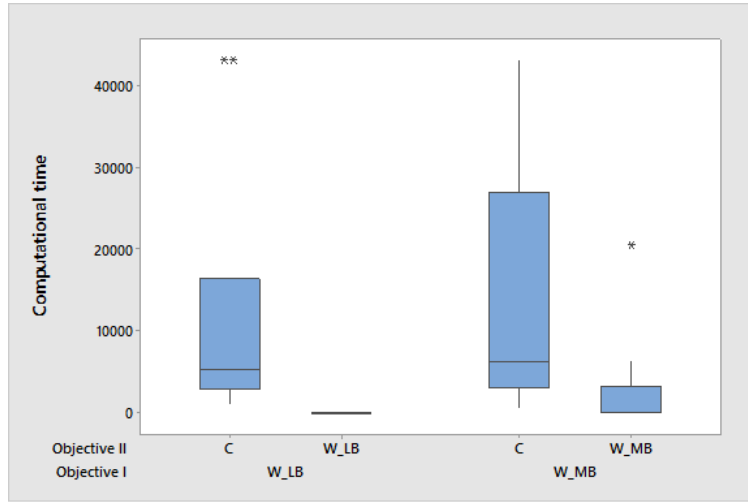
Figure 7: Efficiency analysis of $A + S|S + R$: computational time in seconds (medium-scale instances)

more than the 50% of the experiments, in 1 hour of computation. Furthermore, there are no meaningful improvements in extending the computational time to 6 hours. See Figure 13 for more details. Therefore, such a pattern-based decomposition approach appears to be very effective in computing, in reasonable amounts of time, very good solutions in terms of percentage gap.

5.3 Quality analysis

In this section, we analyze the impact of the pattern-based decomposition on the quality of the computed solutions. The indicators *total travel time* (i.e., the overall distance traveled by the operators) and *balance* (i.e., the difference between the utilization factor of the most and the least busy operators) are considered.

Figures 14 and 15 show the boxplots of the balance and of the total travel time indicators for the small-scale instances, respectively.

The single-phase method was able to determine solutions of good quality concerning the balance criterion W_{LB}. On the other hand, the computed solutions show a large total travel time. This is due to the fact that a very large solution space is explored, thus the single-phase method may incur into a big effort to find near optimal solutions, especially when a cost criterion is adopted.

Different is the behavior of the two-phase decomposition methods.

Minimizing costs in the second phase leads to larger balancing but smaller total travel time than balancing in the second phase. In this case the two-phase methods perform almost similarly concerning the balance criterion, but $A + S|S + R$ appears to be preferable regarding the total travel time, especially when the

31
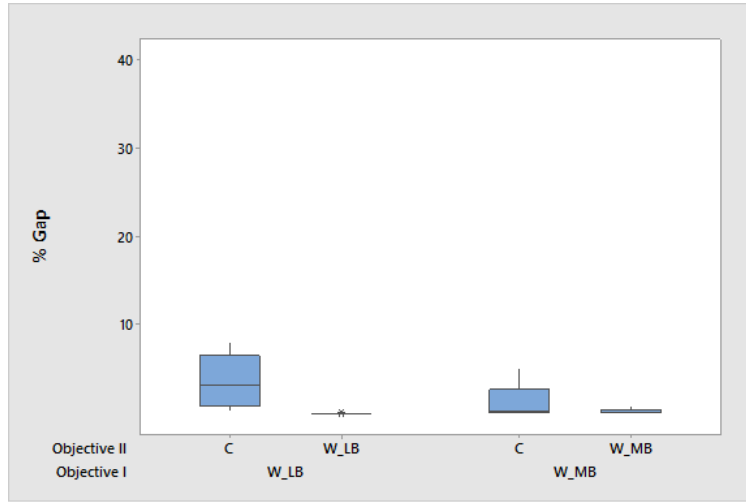
Figure 8: Efficiency analysis of $A + S|S + R$: gap (medium-scale instances)

cost criterion is combined with W_{LB}. In any case, the two-phase decomposition provided much better solutions than the single-phase method by considering the total travel time. It is worthy to point out that, in this scenario, the single-phase method is guided only by a cost criterion and it ignores the balancing issues. In contrast, the two-phases methods fully get the opportunity, given by the problem decomposition, of controlling both criteria in the two phases. See also Figure 16 for an overview of the method behavior in terms of the balance and of the total travel time indicators.

When the objective in the second phase is balancing, all the methods behave similarly in terms of balancing, with a better behavior $A + S|S + R$ when using W_{MB}. See also Figure 17.

By summarizing, the pattern-based decomposition methods allow generally to obtain better quality solutions than the single-phase method, especially regarding the total travel time. This is especially true when controlling the cost in the second phase, although a larger computational time often incurres.

Furthermore, the best combination for the small-scale instances seems to be the $A + S|S + R$ method, as emerged from the analysis of efficiency in Section 5.2. In particular, in order to take the total travel time under control, the version using the cost criterion C in the second phase appears to be preferable, while there is no strong evidence to privilege W_{LB} or W_{MB} in the first phase.

This trend is confirmed when analysing the larger instances. As an example, Figure 18 shows the balance and the total travel time of the medium and of the large-scale solutions computed by $A + S|S + R$ with W_{LB} in the first phase and W_{LB} or C in the second phase. The variant using the criterion C in the second phase
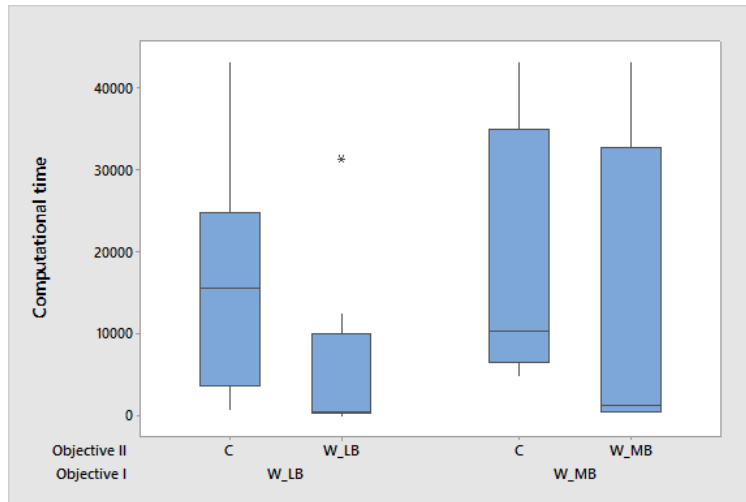
32

Figure 9: Efficiency analysis of $A+S|S+R$: computational time in seconds (large-scale instances)

appears to dominate the others especially in terms of total travel time.

Lastly, since the presented results are based on the hierarchical skill management policy, we provide some information on the number of overskilled visits, i.e. the requests of skill 1 patients that have been served by operators of skill 2. Figure 19 shows this indicator for the small-scale instances, by considering the W_{LB} and the W_{MB} optimization criteria. It can be observed that the *overskilling* is exploited by all the methods, and so this seems to be a relevant option to efficiently address problems arising in HHC.

6   Conclusion

This paper has investigated on four methods, with different level of flexibility, to solve the assignment, scheduling and routing problems of operators in home health care services characterized by skill qualifications, continuity of care and multi-period planning horizon. On the basis of numerical experiments we can draw the following conclusions:

- The most flexible method, i.e. the single-phase one, is affordable only for small-size instances and only when a balancing criterion is considered. As the dimension of the instances increases or cost matters, the computational time required to get a good solution can be quite high.
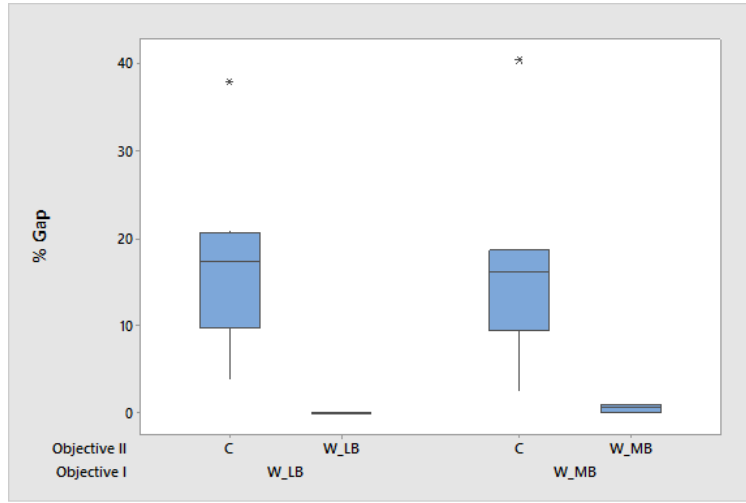
Figure 10: Efficiency analysis of $A + S|S + R$: gap (large-scale instances)

- The most rigid method, in which the first phase is guided only by assignment decisions, failed to provide a feasible solution, at least on the test bed considered, whatever the criteria used to drive the two phases, whatever the skill management strategy adopted.

- The two-phase decomposition method $A + S|R$ and particularly the method $A + S|S + R$ are able to conjugate computational efficiency and goodness of the solutions provided in terms of balance between operators' workload and total travel time.

- The method $A + S|S + R$ is efficient and robust with respect to the optimization criteria used in the two phases. Specifically, for all the instances in the test bed, whatever the dimension, when used in combination with the balancing criterion W_{LB} in both phases, it is very efficient and returns, in most cases, solutions with a very low percentage gap. In addition, the versions of $A + S|S + R$ using the cost criterion in the second phase, is very robust, being capable to compute a feasible solution in all the tested cases, although they required an increased computational time with respect to the variants characterized only by a balance criterion. The method $A + S|S + R$ is also capable to compute near optimal solutions even when a short time limit is imposed; indeed, it is able to provide a solution within 5% from the best bound in 15 minutes for more than 50% of the runs; in addition, no meaningful improvements are obtained further extending the time limit.

Summarizing, the main achievement of the computational campaign is that the proposed pattern-based decomposition $A + S|S + R$ is a valuable approach to
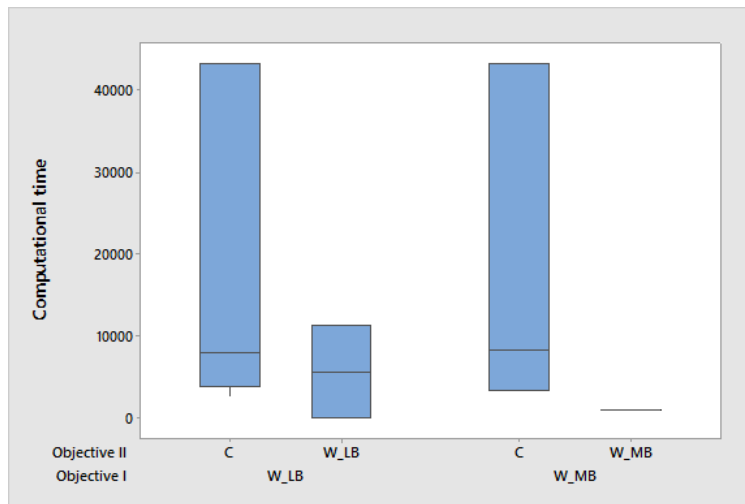
Figure 11: Efficiency analysis of $A+S|S+R$: computational time in seconds (big instances)

improve on the computational efficiency of the single-phase method, without substantially deteriorating its flexibility, i.e. the quality of the returned solutions. Clearly, the above conclusions are limited to the numerical experiments reported in this paper. However, they encourage further work towards the design of more sophisticated exact and heuristic algorithms, relying on the proposed pattern decomposition. Other research developments could be devoted to enrich the setting addressed by considering time windows and synchronization constraints into the problem as well as multi-objective functions.

References

Akjiratikarl, C., P. Yenradee, P. R. Drake. 2007. Pso-based algorithm for home care worker scheduling in the united kingdom. *Computers & Industrial Engineering* **53**(4) 559 – 583.

Begur, S.V., D.M. Miller, J.R. Weaver. 1997. An integrated spatial dss for scheduling and routing home-health-care nurses. *Interfaces* **27**(4) 35–48.

Borsani, V., A. Matta, G. Beschi, F. Sommaruga. 2006. A home care scheduling model for human resources. *Service Systems and Service Management, 2006 International Conference on*, vol. 1. 449–454.

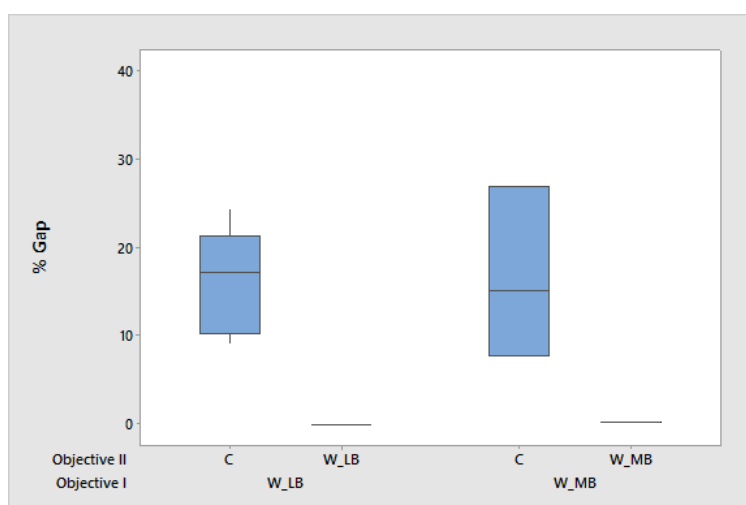Bredstrom, D., M. Ronnqvist. 2007. A branch and price algorithm for the com-

Figure 12: Efficiency analysis of $A + S | S + R$: gap (big instances)

bined vehicle routing and scheduling problem with synchronization constraints. Tech. rep., NHH Dept. of Finance & Management Science.

Bredstrom, D., M. Ronnqvist. 2008. Combined vehicle routing and scheduling with temporal precedence and synchronization constraints. *European Journal of Operational Research* **191**(1) 19 – 31.

Brucker, P., R. Qu, E. Burke. 2011. Personnel scheduling: Models and complexity. *European Journal of Operational Research* **210**(3) 467 – 473.

Cappanera, P., L. Gouveia, M.G. Scutellà. 2013. Models and valid inequalities to asymmetric skill-based routing problems. *EURO Journal on Transportation and Logistics* **2**(1-2) 29–55.

Cappanera, P., M.G. Scutellà. 2013. Home care optimization: impact of pattern generation policies on scheduling and routing decisions. *Electronic Notes in Discrete Mathematics* **41**(0) 53 – 60.

Cappanera, P., M.G. Scutellà. 2015. Joint assignment, scheduling, and routing models to home care optimization: A pattern-based approach. *Transportation Science* **49**(4) 830–852.

Cattafi, M., R. Herrero, M. Gavanelli, M. Nonato, F. Malucelli F. 2012. Improving quality and efficiency in home health care: An application of constraint logic programming for the ferrara nhs unit. Dovier A., V.S. Costa, eds., *Technical Communications of the 28th International Conference on Logic Programming*
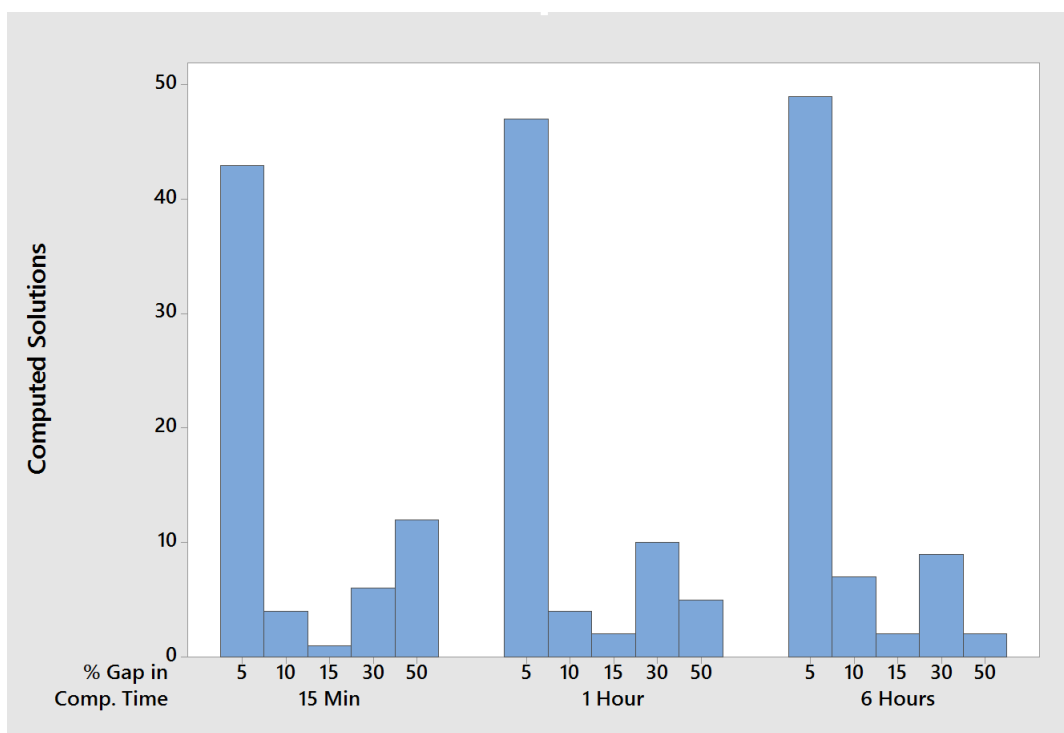
Figure 13: Overall efficiency analysis of $A + S|S + R$: reduced time limit

(ICLP'12), *Leibniz International Proceedings in Informatics (LIPIcs)*, vol. 17. 414–424.

Causmaecker, P. De, P. Demeester, G. Vanden Berghe, B. Verbeke. 2005. Analysis of real-world personnel scheduling problems. *Practice and Theory of Automated Timetabling, International Conference on*. 183–198.

Cheng, E., J.L. Rich. 1998. A home health care routing and scheduling problem.

Eveborn, P., P. Flisberg, M. Ronnqvist. 2006. Laps carean operational system for staff planning of home care. *European Journal of Operational Research* **171**(3) 962 – 976. Feature Cluster: Heuristic and Stochastic Methods in Optimization Feature Cluster: New Opportunities for Operations Research.

Gamst, M., T.Sejr Jensen. 2012. A branch-and-price algorithm for the long-term home care scheduling problem. Diethard Klatte, Hans-Jakob Lthi, Karl Schmedders, eds., *Operations Research Proceedings 2011*. Operations Research Proceedings, Springer Berlin Heidelberg, 483–488.

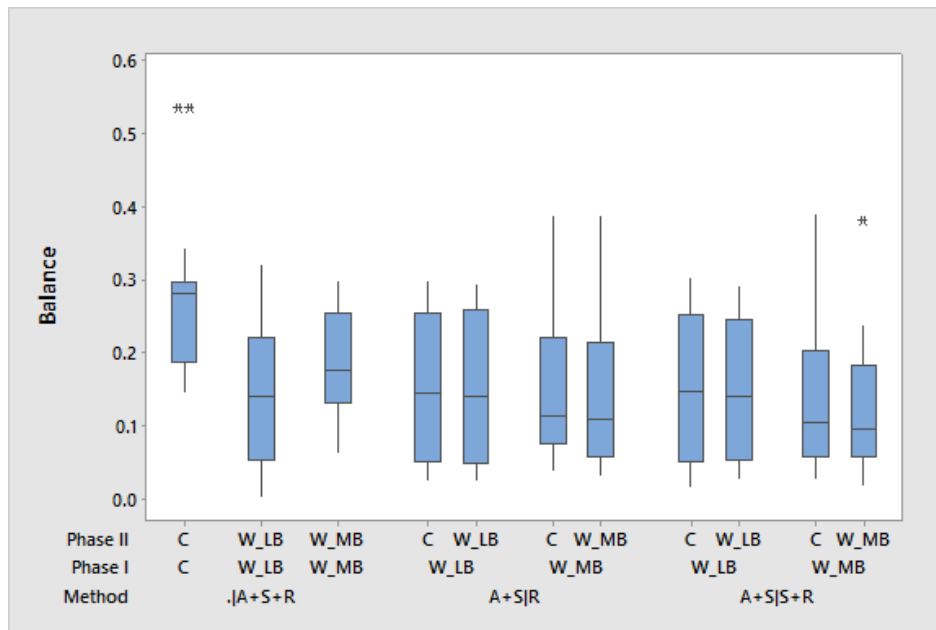Hulshof, P.J.H., N. Kortbeek, R.J. Boucherie, E.W. Hans P.J.M. Bakker. 2012.

Figure 14: Quality analysis of single-phase and two-phase methods: balance (small-scale instances)

Taxonomic classification of planning decisions in health care: a review of the state of the art in or/ms. *Health Systems* **1** 129–175.

Kucukyazici, B., V. Verter, N.E. Mayo. 2011. An analytical framework for designing community-based care for chronic diseases. *Production and Operations Management* **20**(3) 474–488.

Lanzarone, E., A. Matta, E. Sahin. 2012. Operations management applied to home care services: The problem of assigning human resources to patients. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* **42**(6) 1346–1363.

Lanzarone, E., A. Matta, G. Scaccabarozzi. 2010. A patient stochastic model to support human resource planning in home care. *Production Planning and Control* **21**(1) 3–25.

Matta, A., S. Chahed, E. Sahin, Y. Dallery. 2014. Modelling home care organisations from an operations management perspective. *Flexible Services and Manufacturing Journal* **26**(3) 295–319.

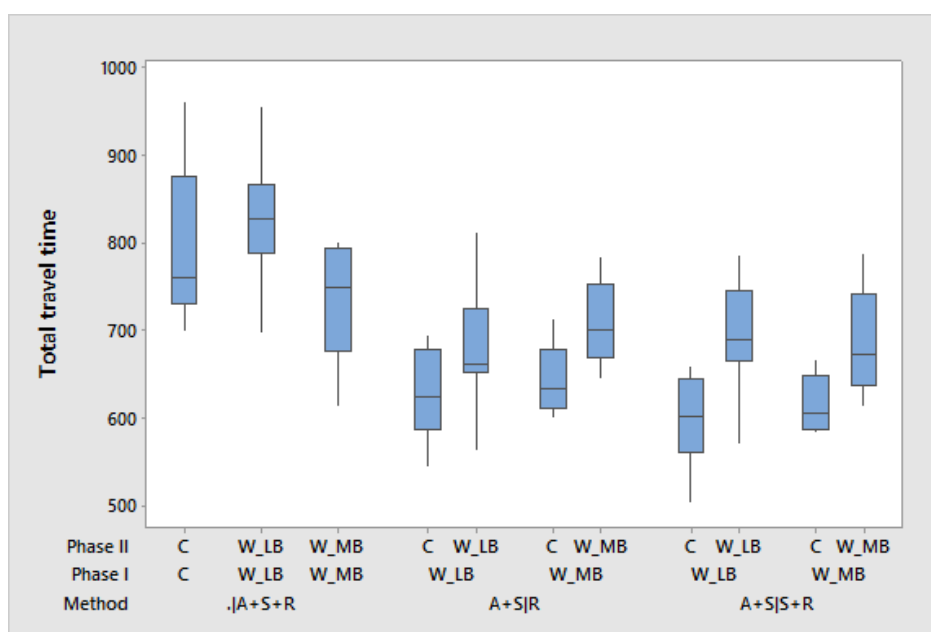Nickel, S., M. Schrder, J. Steeg. 2012. Mid-term and short-term planning sup-

Figure 15: Quality analysis of single-phase and two-phase methods: total travel time (small-scale instances)

port for home health care services. *European Journal of Operational Research* **219**(3) 574 – 587.

Rasmussen, M.S., T. Justesen, A. Dohn, J. Larsen. 2012. The home care crew scheduling problem: Preference-based visit clustering and temporal dependencies. *European Journal of Operational Research* **219**(3) 598 – 610.

Sahin, E., A. Matta. 2014. A contribution to operations management-related issues and models for home care structures. *International Journal of Logistics Research and Applications* **0**(0) 1–31.

Thomsen, K. 2006. Optimization on home care. URL http://www2.imm.dtu.dk/pubdb/p.php?4710. Supervised Jesper Larsen, IMM, DTU, and co-supervisor Rene Munk Jørgensen, CTT, DTU.

Trautsamwieser, A., M. Gronalt, P. Hirsch. 2011. Securing home health care in times of natural disasters. *OR Spectrum* **33**(3) 787–813.

Trautsamwieser, Andrea, Patrick Hirsch. 2014. A branch-price-and-cut approach for solving the medium-term home health care planning problem. *Networks* **64**(3) 143–159.
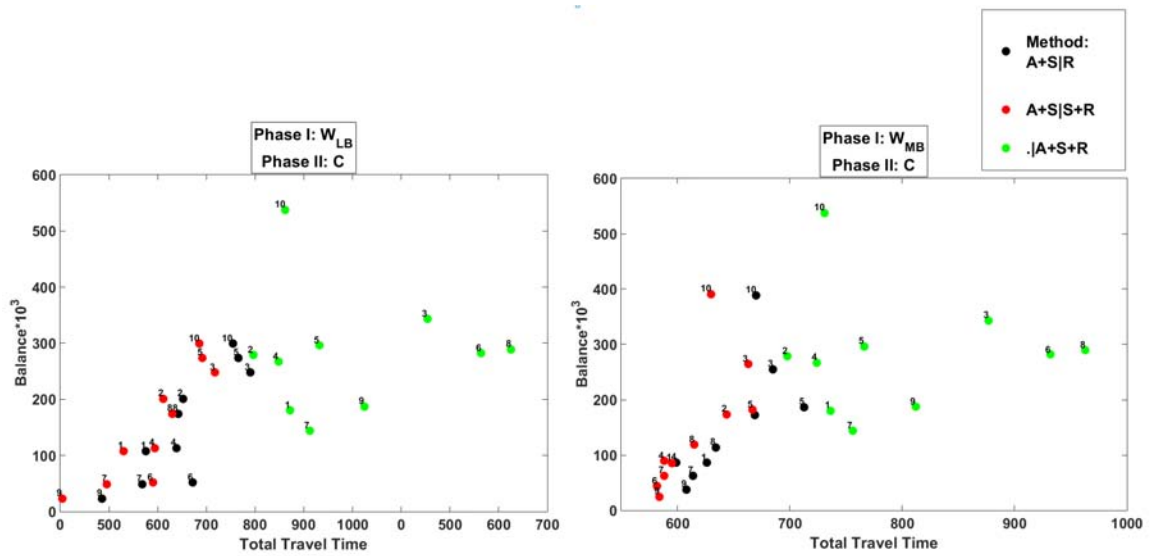
Figure 16: Cost criterion in the second phase (small-case instances)

Yalcindag, S., A. Matta, E. Sahin. 2012. Human resource scheduling and routing problems in home health care context: A literature review. Tech. rep., Laboratoire Gnie Industriel, Ecole Centrale Paris. URL http://www.lgi.ecp.fr/Biblio/PDF/CR-LGI-2012-10.pdf. Cahiers de recherche 2012-10.
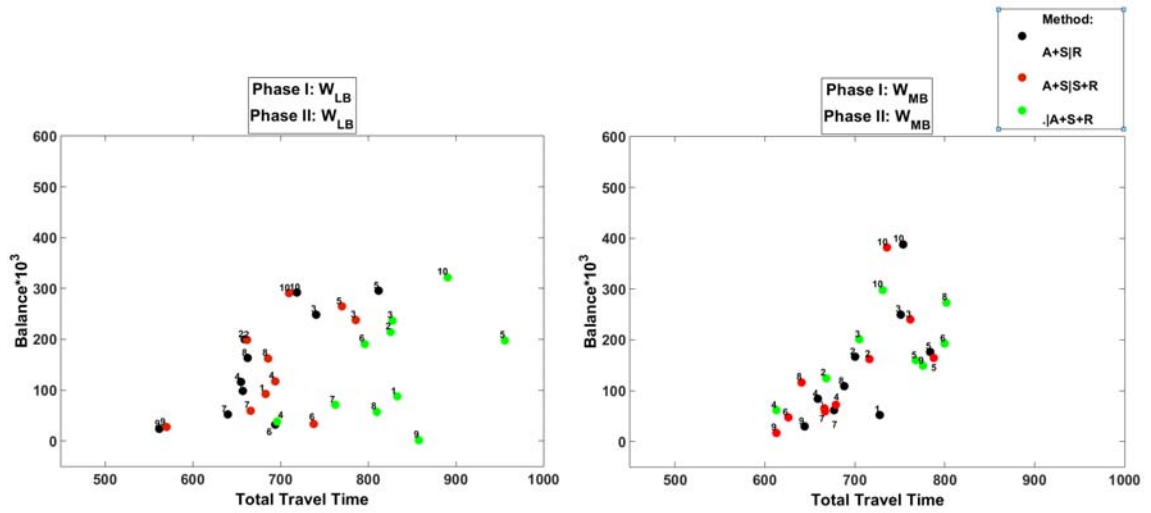
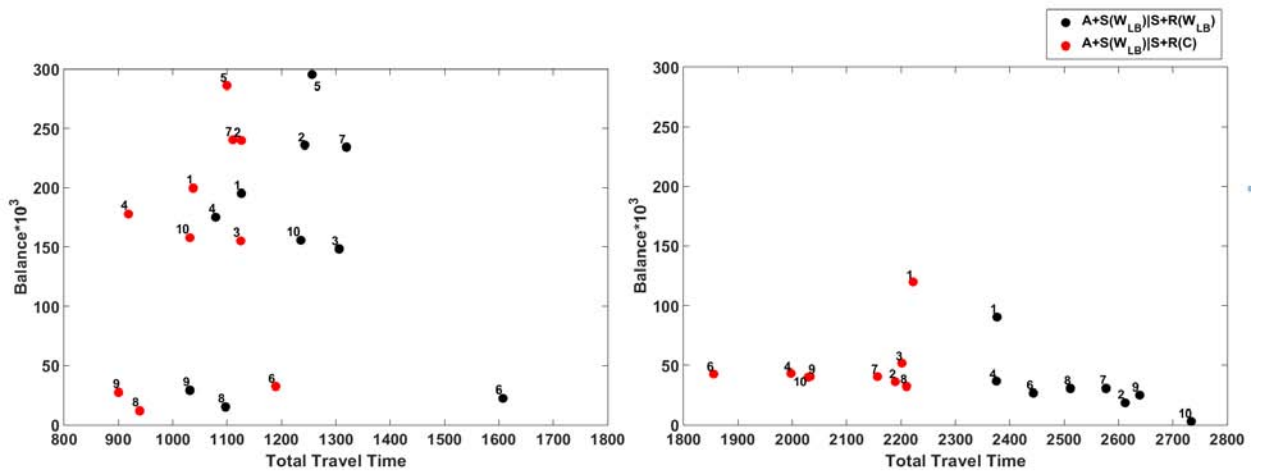Figure 17: Balancing criterion in the second phase (small-case instances)

Figure 18: Quality analysis of $A+S|S+R$ (medium-scale instances on the left and large-scale instances on the right)
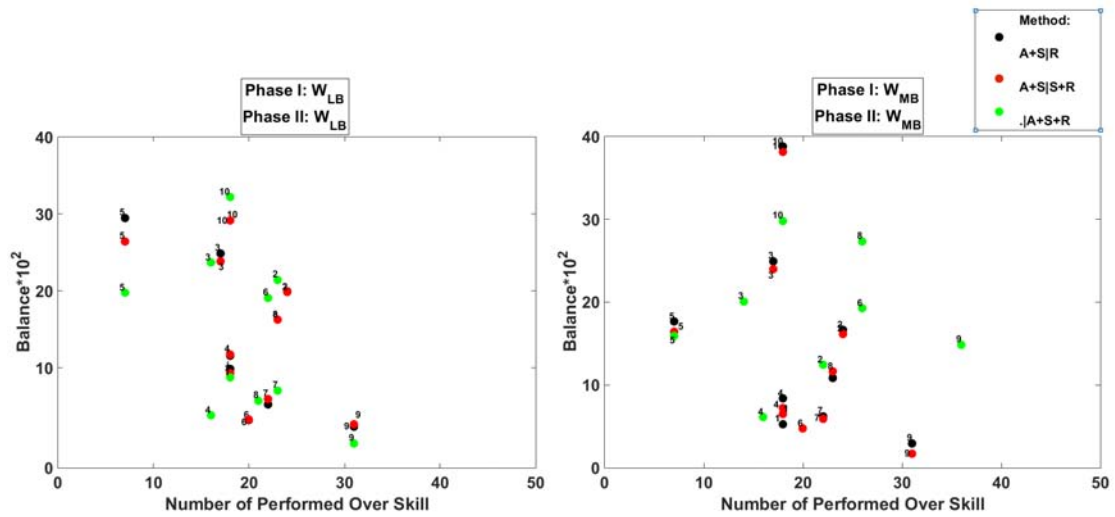
41

Figure 19: Overskill analysis (small-scale instances)